A versatile dual reporter to identify ribosome pausing motifs alleviated by translation elongation factor P

Urte Tomasiunaite [1], Tess Brewer [1], Korinna Burdack [1], Sophie Brameyer [1], Kirsten Jung [1*]

[1] Faculty of Biology, Microbiology, Ludwig-Maximilians-Universität München, Martinsried, Germany

*Correspondence: jung@lmu.de

Keywords: stalling motifs; ribosome stalling; screening for ribosome pausing strength; translational efficiency; dual-reporter; proteins with polyproline motifs; elongation factor P

Abstract

Protein synthesis is influenced by the chemical and structural properties of the amino acids incorporated into the polypeptide chain. Motifs with consecutive prolines can slow down translation speed and cause ribosome stalling. Translation elongation factor P (EF-P) facilitates peptide bond formation in these motifs, thereby alleviating stalled ribosomes and restoring regular translational speed. Ribosome pausing at various polyproline motifs has been intensively studied using a range of sophisticated techniques, including ribosome profiling, proteomics, and *in vivo* screenings with reporters incorporated into the chromosome. However, the full spectrum of motifs which cause translational pausing in *Escherichia coli* has not yet been identified. Here we describe a plasmid-based dual reporter for rapid assessment of pausing motifs. This reporter contains two coupled genes encoding mScarlet-I and chloramphenicol acetyltransferase to screen motif libraries based on both bacterial fluorescence and survival. In combination with a diprolyl motif library, we use this reporter to reveal motifs of different pausing strengths in an *E. coli* strain lacking *efp.* Subsequently, we use the reporter for a high-throughput screen of four motif libraries, with and without prolines at different positions, sorted by fluorescence-associated cell sorting (FACS) and identify new motifs that influence translational efficiency of the fluorophore. Our study provides an *in vivo* platform for rapid screening of amino acid motifs that affect translational efficiencies.

Introduction

The regulation of protein synthesis is critical for the maintenance of physiological balance and the survival of organisms. A number of factors can modulate translational efficiency, including charge-specific interactions between the ribosome peptide exit tunnel and the nascent chain[1-3], mRNA secondary structures[4-7], tRNA abundance and modification[8-10], and codon selection[11,12]. It has

been widely demonstrated that certain combinations of amino acids, in particular stretches of polyprolines, have the ability to negatively impact translational efficiency and in certain instances cause the ribosome to stall[13-16]. The profound negative effect of proline on translation rate can be attributed to its unique structural nature, which includes a pyrrolidine ring that renders it conformationally rigid. This ultimately results in proline being a poor donor and poor acceptor during the transpeptidation reaction[17-20]. Despite the fact that proline is a challenging factor during translation, diprolyl motifs (XPPX) are very common in bacterial proteomes. In *Escherichia coli*, there are 2,184 of these motifs (0.51 motif/protein), in *Salmonella enterica* 2,324 motifs (0.51 motif/protein), in *Pseudomonas aeruginosa* 4,074 motifs (0.73 motif/protein) and *Streptomyces coelicolor* even contains 8,719 motifs (1.08 motif/protein)[21,22]. The prevalence of these motifs is not limited to bacterial proteomes — they are also common among the proteomes of eukaryotes[23]. The ubiquity of these motifs in the proteomes of organisms underlines their functional and regulatory importance to shape protein structure[24,25], adjust copy number[21] and incorporate transmembrane proteins[26].

The biosynthesis of proteins containing polyproline motifs is mediated by distinct translation factors that have evolved in organisms from all domains of life. While eukaryotes and archaea encode initiation factor 5A (e/aIF5A)[27], bacteria have evolved elongation factor P (EF-P), which reduces ribosomal stalling by assisting during the translation of polyproline motifs[14,15]. The EF-Ps of numerous bacterial species are dependent on unique post-translational modifications (PTM), which are necessary to improve their activity[28-35]. In contrast, certain EF-Ps, particularly from the subfamily with the activation loop PGKGP (Pro-Gly-Lys-Gly-Pro), exhibit full functionality in the absence of PTMs, although with reduced levels of activity in *E. coli*[22,36]. Intriguingly, our recent findings even demonstrate that some unmodified EF-Ps, which are barely active in *E. coli*, can be engineered functional in *E. coli*[36]. This suggests that full EF-P functionality is not exclusively dependent on PTM, as previously assumed. Notably, we observed variations in the efficacy of distinct polyproline motif rescue among modified and unmodified EF-Ps, suggesting that the underlying general mechanisms of protein synthesis are not yet fully understood and require further investigation.

A number of studies have employed a range of techniques, including ribosome profiling, mass spectrometry, in *vitro* and in *vivo* reporter systems, to identify motifs associated with different degrees of ribosome pausing[14,16,37-41]. The majority of these studies focuses on motifs which contain at least two prolines, as EF-P is well known to support their synthesis during the transpeptidation reaction[16,38]. The literature on this topic is limited with regards to motifs that contain a single proline or even no proline, which rescue is dependent on a functional EF-P[37,42]. The existing knowledge gaps highlight the need for simple, accessible tools that can be utilized by every laboratory at a low cost, with a minimal input of materials to investigate translational dynamics.

In this study, we present a user-friendly, plasmid-based reporter system for investigating translational efficiencies of amino acid motifs. Translationally coupling[43] the genes encoding the fluorophore mScarlet-I and chloramphenicol acetyltransferase, enhances screening capacity, allowing dual output measurements of cell fluorescence and viability. High-throughput screening demonstrates that this system effectively identifies unusual sequence motifs, particularly those lacking proline, which cause reduced translation rates.

Results

A plasmid-based in *vivo* dual reporter system enables ribosome pausing measurements in a Δ*efp* mutant

To facilitate rapid evaluation of which amino acid motifs can influence translational efficiency, we established a plasmid-based reporter system. Motifs which are known to cause various levels of ribosome pausing were chosen for initial experiments: RPAP (Arg-Pro-Ala-Pro) – represented no pausing strength, TPPP (Thr-Pro-Pro-Pro) – represented middle pausing strength, and RPPP (Arg-Pro-Pro-Pro) – represented strong pausing strength[14,38,41] (**Figure 1A**). To identify differences in translational efficiency resulting from distinct sequence motifs, an *E. coli* BW25113 mutant lacking *elongation factor P* (Δ*efp*) was transformed with these reporter plasmids. The initial reporter design consisted of an antibiotic resistance cassette (*chloramphenicol acetyltransferase, cat*) with motifs integrated at the N-terminus (**Supplementary Figure 1A,** left panel). This approach allowed estimation of the translational efficiencies of *cat* influenced by the motif sequence, based on bacterial survival in the presence of chloramphenicol. A 24-hour incubation in media containing serial dilutions of chloramphenicol demonstrated differences in bacterial survival only between clones carrying the reporters RPAP_*cat* and RPPP_*cat,* particularly at chloramphenicol concentrations 3.4 µg/mL, 3.8 µg/mL and 4.3 µg/mL (**Supplementary Figure 1B-C**). However, no significant differences between the clones carrying the reporters RPAP_*cat* and TPPP_*cat* could be detected at any of the chloramphenicol concentrations tested (**Supplementary Figure 1C**).

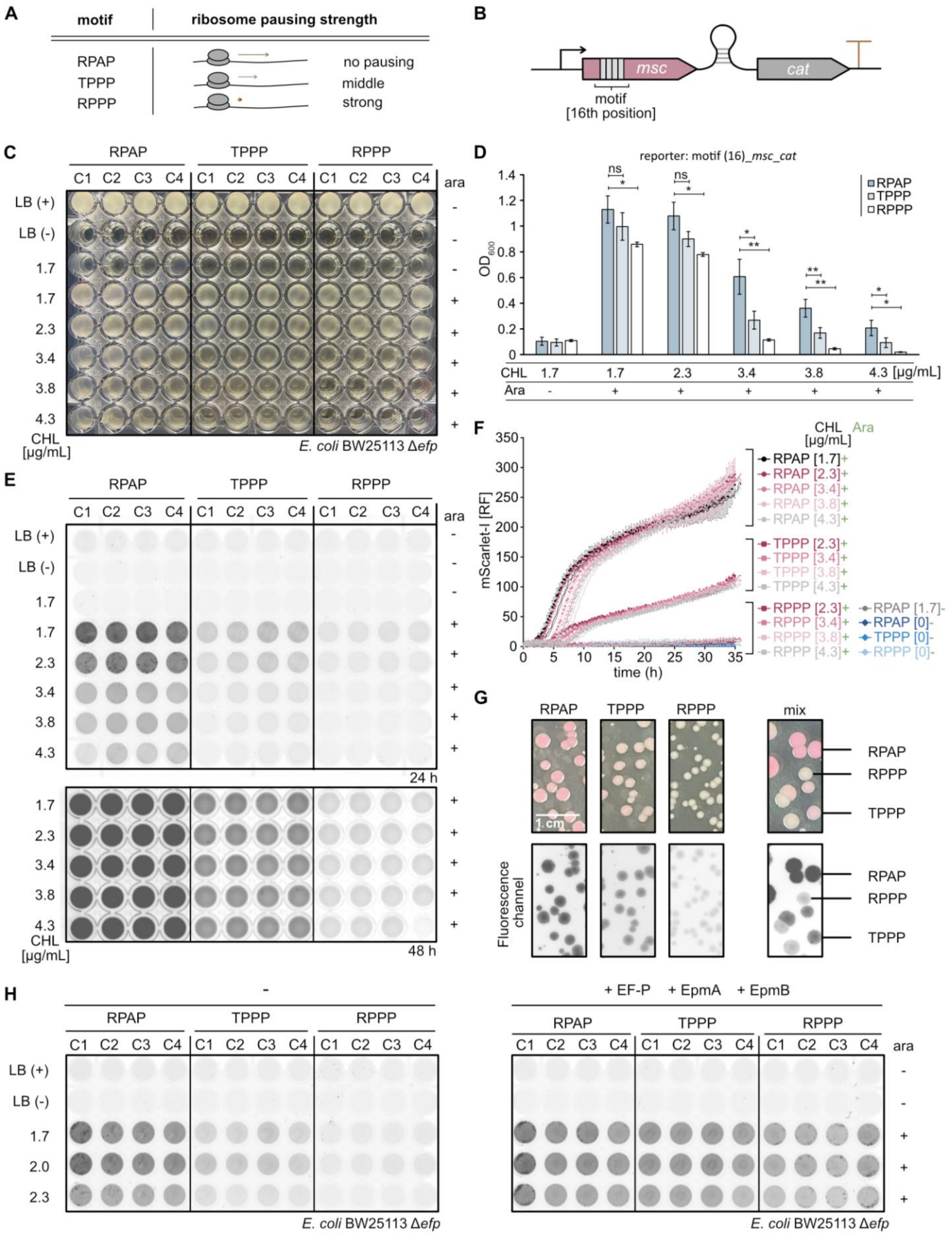
We hypothesized that reducing *cat* expression would enhance screening sensitivity, making the translational efficiencies of the RPAP and TPPP motifs more distinguishable. As previously described[43,44], a weak translational coupling of a fluorophore gene upstream of a resistance gene cassette can support attenuation of antibiotic resistance and enable fine-tuning of screenings. Consequently, we constructed a dual reporter plasmid encoding for super folded green fluorescent protein (sfGFP) and CAT with a weak translational coupling device[43] connecting the genes (**Supplementary Figure 1A**, middle and right panel). Motifs were incorporated into the gene encoding the fluorophore (**Supplementary Figure 1A**, middle and right panel). Endpoint optical density quantification revealed that chloramphenicol influenced the survival of all clones tested with the dual reporter motif_*sfgfp_cat*, regardless of the specific motif present

(**Supplementary Figure 1D**). Significant differences between all clones in media supplemented with 3.4 µg/mL or 3.8 µg/mL chloramphenicol were only detectable when motifs were at the sixteenth nucleotide position of the fluorophore sequence (**Supplementary Figure 1E**). This observation confirms the essentiality of the first fifteen nucleotides which are critical for the initial synthesis of the nascent peptide chain[45].

To increase visibility on agar plates, we replaced *sfgfp* with *mscarlet-I*, since the fluorophore mScarlet-I changes the colony color on agar plates to bright pink as shown in earlier studies[46] (**Figure 1B**). Endpoint quantification of optical densities demonstrated that all Δ*efp* cells containing reporters encoding RPAP, TPPP and RPPP motifs exhibited differences in survival at 3.4 µg/mL, 3.8 µg/mL and 4.3 µg/mL chloramphenicol concentrations (**Figure 1C and D**). In particular, cells carrying the reporter RPPP_*mscarlett-I_cat* struggled to survive at chloramphenicol concentrations of 3.4 µg/mL or higher (**Figure 1C and D**). These findings imply that the translational efficiency of different motifs can be measured with the dual reporter system based on bacterial survival. Fluorescence intensity measurements after 24 hours (**Figure 1E**, upper panel) were in line with the observations from the bacterial survival assay (**Figure 1C-D**). We also extended the incubation period to 48-hours to ensure an equal optical density of bacterial culture in all wells, thereby demonstrating that the low fluorescence intensities observed with the RPPP_*mscarlett-I_cat* (**Figure 1E**, upper panel) reporter were not caused by differential cell abundance (**Figure 1E**, bottom panel).

We monitored relative fluorescence intensities for 35 hours in M9 minimal medium to investigate translational efficiencies over time (**Figure 1F**). The Δ*efp* strain with the reporter RPAP_*mscarlett-I_cat* exhibited the greatest relative fluorescence (RF) intensities, reaching a range of 200-300 RF between 25 and 35 hours (**Figure 1F**). The initial ten hours of monitoring revealed the impact of the supplemented chloramphenicol concentrations on the relative fluorescence (**Figure 1F**). The Δ*efp* strain containing the reporter TPPP_*mscarlett-I_cat* exhibited intermediate fluorescence intensities, reaching a range of 80-110 RFU between 25 and 35 hours (**Figure 1F**). Here, the impact of the supplemented chloramphenicol concentrations on the relative fluorescence curves was detected at the time interval of four to fourteen hours (**Figure 1F**). The low relative fluorescence intensities observed with strains containing the reporter plasmid RPPP_*mscarlett-I_cat,* were consistent with the endpoint fluorescence intensity measurements in **Figure 1E**. Regardless of the motif present, fluorescence was consistently low in the absence of arabinose and chloramphenicol, indicating low levels of leaky expression (**Figure 1F).**

To assess the reporter's suitability for plate-based screens, we examined the growth of the Δ*efp* strain carrying reporters with RPAP, TPPP, or RPPP motifs on LB agar plates (**Figure 1G**). Preliminary analysis showed that LB agar plates with 3.4 µg/ mL chloramphenicol exhibited the most significant differences in colony morphology and fluorescence between cells containing reporters with either the RPAP or RPPP motifs (**Supplementary Figure 2A-C**).



139 **Figure 1. Development of a dual reporter system to screen for ribosome pausing motifs.**

140 **(*see description on the next page*)**

Figure 1 (description). **A** Schematic of known ribosome pausing strengths at motifs RPAP (Arg-Pro-Ala-Pro), TPPP (Thr-Pro-Pro-Pro) and RPPP (Arg-Pro-Pro-Pro) alleviated by elongation factor P[14,38,41]. **B** Construction of dual reporter for translational efficiency measurements. The reporter consists of a gene coding for fluorophore (mScarlet-I) and a resistance gene cassette (*chloramphenicol acetyltransferase, cat*), both linked by a weak translational coupling device[43]. **C-D** 96-well assay plate of the growth measurements (**C**) and growth quantification (**D**) of *E. coli* BW25113 Δ*efp*, transformed with the dual reporter plasmid pBAD24_motif_*mscarlet-I_cat*. **E** Fluorescence measurements of the 96-well assay plate 24 hours (upper panel) or 48 hours (bottom panel) after bacteria inoculation. *E. coli* BW25113 Δ*efp* was transformed with the dual reporter plasmid pBAD24_motif_*mscarlet-I_cat*. **F** Relative fluorescence (RF) measurements over time of *E. coli* BW25113 Δ*efp* with dual reporter plasmid pBAD24_motif_*mscarlet-I_cat*. The M9 medium was supplemented with different concentrations of chloramphenicol (CHL). mScarlet-I fluorescence signal was normalized to the measured optical density (OD_{600}) and shown as RF. **G** Colony morphology and fluorescence analysis of *E. coli* BW25113 Δ*efp* with dual reporter plasmid pBAD24_motif_*mscarlet-I_cat*. Bacteria were plated on LB agar assay plates supplemented with 0.2 % (w/v) arabinose (Ara) and 3.4 µg/mL CHL. Nucleotides encoding the RPAP, TPPP, and RPPP motifs were incorporated into the fluorophore gene sequence starting from the sixteenth nucleotide position (16th). **H** EF-P complementation assay. Fluorescence measurements of the 96-well assay plate 30 hours after bacteria inoculation. *E. coli* BW25113 Δ*efp* was transformed with the dual reporter plasmid pBAD24_motif_*mscarlet-I_cat* (left panel) or with the dual reporter plasmid pBAD24_motif_*mscarlet-I_cat* and pBAD_*E. coli efp_epmA_epmB* (right panel). Error bars indicate the standard deviation (SD) of three independent biological replicates. Statistics: student's unpaired two-sided t test (****p < 0,0001; ***p < 0.001; **p < 0.01; *p < 0.05; ns p > 0.05). 1.7 µg/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: ns p = 0.178 [D]; RPAP vs RPPP: *p = 0.016 [D]); 2.3 µg/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: ns p = 0.055 [D]; RPAP vs RPPP: *p = 0.013 [D]); 3.4 µg/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: *p = 0.015 [D]; RPAP vs RPPP: **p = 0.008 [D]); 3.8 µg/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: **p = 0.009 [D]; RPAP vs RPPP: **p = 0.003 [D]); 4.3 µg/mL CHL and 0.2 % (w/v) Ara (RPAP vs TPPP: *p = 0.038 [D]; RPAP vs RPPP: *p = 0.012 [D]). C1-C4 – clone number.

After a 120-hour incubation period at room temperature, the colonies of Δ*efp* with RPAP_*mscarlett-I_cat* exhibited pink color and the largest colony diameter (**Figure 1G**). In contrast, Δ*efp* colonies expressing RPPP_*mscarlett-I_cat* exhibited yellow-white color and smaller colony diameters. In line with our measurements, Δ*efp* with TPPP_*mscarlett-I_cat* exhibited a light pink color and an intermediate colony diameter (**Figure 1G**). When EF-P and its modification machinery were overproduced in the Δ*efp* strain, translation of the fluorophore was rescued for all motifs (**Figure 1H**). This finding suggests that the observed reduction in fluorescence in the RPPP and TPPP reporters in Δ*efp* can be attributed to ribosome pausing events caused by the motifs in the absence of EF-P.

Taken together, our results demonstrate that the strength of ribosome pausing at amino acid motifs, can be effectively measured using a dual reporter system. Overall, these data show that mScarlet-I is an appropriate choice for the dual reporter system, enabling sensitive colony color contrast- and fluorescence-based screenings on plates to assess ribosome pausing strengths at specific motifs.

Screening of libraries with diprolyl motifs demonstrates variation in translational efficiencies

We used plate-based screening of a diprolyl library (XPPX, X-Pro-Pro-X) in Δ*efp* to assess the efficacy of our dual reporter for large-scale translational efficiency measurements. We used a three-step protocol to generate the XPPX library, as illustrated in detail in **Figure 2A**. First, the dual reporter plasmid was amplified using primers with complementary overhangs. The forward primer was designed to be degenerate, incorporating all possible nucleotide (N) combinations between the two CCG codons coding for prolines (NNN-CCG-CCG-NNN) in the motif (**Figure 2A**). This strategy enables the incorporation of any amino acid in the positions upstream and downstream of the two prolines in the motif XPPX. Second, the *E. coli* MC1061 strain was transformed with the linearized XPPX library plasmids, allowing the library to circularize[43,44,47]. Third, the circularized XPPX library was isolated and the Δ*efp* mutant was transformed with the library (**Figure 2A**). The bacterial library populations exhibited diverse colony morphologies after plating on assay plates (**Figure 2B**). Qualitatively, we found that the colonies with the biggest diameter had a deep pink color, while the smaller colonies tended to be lighter pink or white-yellow (**Figure 2B**). The fluorescence analysis confirmed our previous observation that increasing levels of pink pigment resulted in higher fluorescence (**Figure 1G** and **Figure 2B**). These observations confirm that the amino acids which surround the two prolines in a diprolyl motif determine its translational efficiency, in line with published studies[16,21,39].

As a first pass, we manually screened sixty clones by grouping them into three categories based on colony color (pink, light pink and white-yellow colonies) and determined their motifs by Sanger sequencing (**Figure 2C**). We found some motifs overlapped in adjacent categories (for example between pink and light pink colonies) but found no overlaps between the pink and white-yellow color categories. Overlapping motifs were re-categorized into the category associated with weaker effects on translational efficiency. Altogether, we identified twenty-six sequences which had a marginal impact on the translational efficiency of the fluorophore (colonies with pink color), twenty-two sequences which caused moderate decreases in translational efficiency (colonies with light pink color), and twelve sequences which caused a strong decrease in translational efficiency (colonies with white-yellow color) (**Figure 2C**). Interestingly, motifs encoding FPPX were found across all categories of translational efficiency (X indicates any amino acid). When encoded in the FPPX motif, phenylalanine, isoleucine, valine and tyrosine (FPPF, FPPI, FPPV and FPPY) had a marginal influence on translational efficiency of the fluorophore. Cysteine, glycine and arginine (FPPC, FPPG and FPPR) caused a moderate decrease in translational efficiency, and proline (FPPP) had the strongest effect. These results indicate that the amino acid downstream of the two prolines in a diprolyl motif critically determines translational efficiency, consistent with previous findings[21].

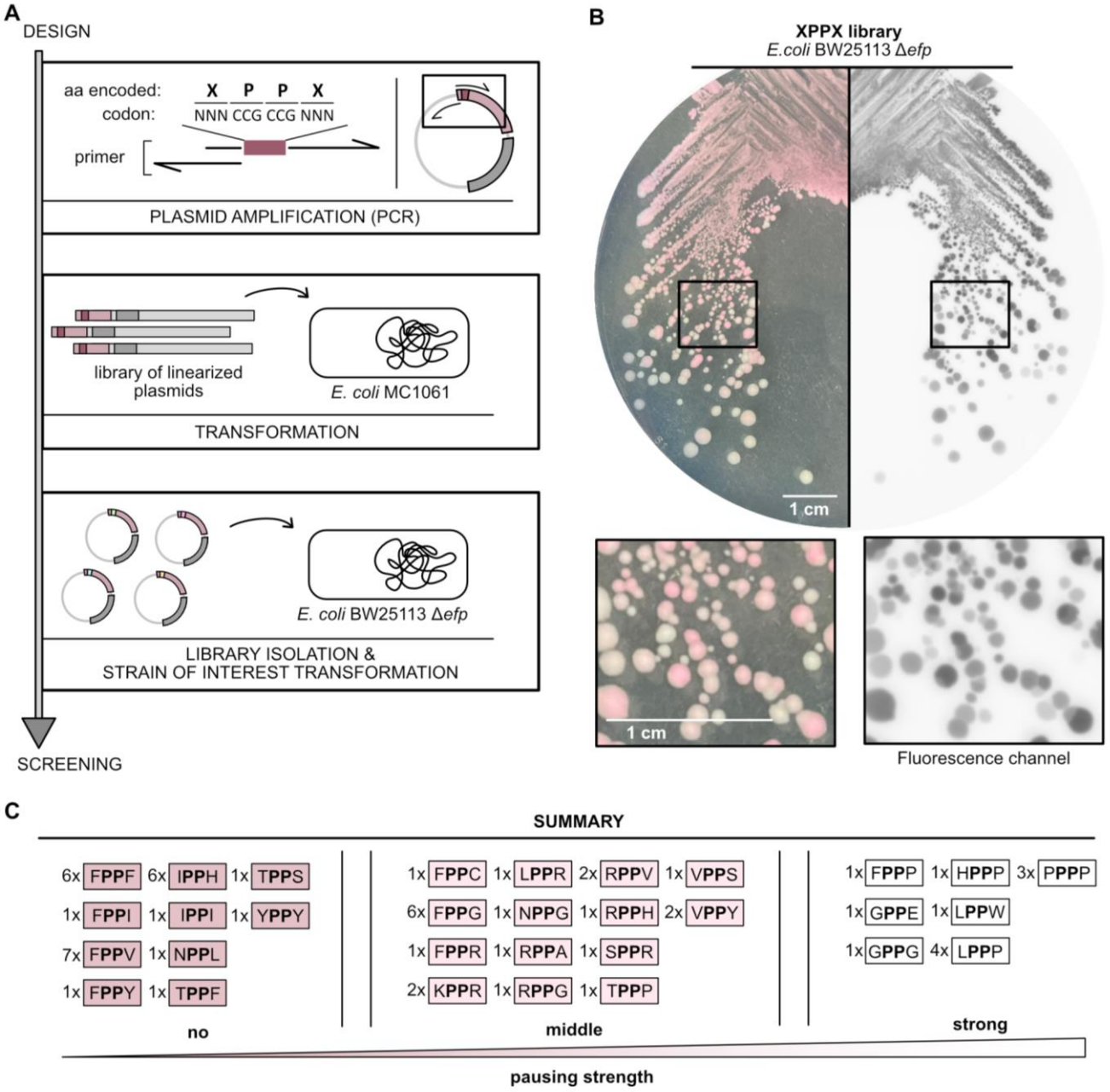


Figure 2. XPPX library screening on agar assay plates distinguishes motifs with different pausing strengths. A Graphical workflow to generate the XPPX (X-Pro-Pro-X) library (X represents all possible amino acids, N – all possible nucleotides). **B** Colony morphology and fluorescence analysis of *E. coli* BW25113 Δ*efp*, transformed with dual reporter plasmid pBAD24_XPPX_*mscarlet-I_cat* encoding the XPPX library. The images in the rectangles are shown enlarged under the agar plate. **C** Summary of the sequencing result of 60 clones that were picked according to their fluorescence intensity and viability. Motifs identified in pink colonies with high fluorescence intensities and large colony diameters were classified as those causing no pausing strength. Motifs observed in light pink colonies with intermediate fluorescence intensities and intermediate colony diameter were categorised as those causing middle pausing strength. Motifs identified in white-yellow colonies with low fluorescence intensities and small colony diameters were grouped as those causing strong pausing strength. The numbers indicate the count of plasmids found, that contain the indicated motif. aa – amino acid.

Together, these results demonstrate that our dual reporter can differentiate the pausing strength of diprolyl motifs in a Δ*efp* strain. We show that the strength of ribosome pausing can be assessed on agar plates supplemented with chloramphenicol by analyzing viability and the fluorescence of the colonies.

High throughput screening uncovers new motifs that reduce the translational efficiencies

To assess the reporter's capability for high throughput translational efficiency measurements, we expanded our investigations using flow cytometry. To enhance the diversity of motifs we could analyze, we designed additional libraries with either a fixed CCG codon coding for a single proline (XPXX, NNN-CCG-NNN-NNN; XXPX, NNN-NNN-CCG-NNN) or completely randomized codons (XXXX, NNN-NNN-NNN-NNN). We sorted cells containing these plasmid libraries (about one million cells per library) with fluorescence-associated cell sorting (FACS) and analyzed the sorted cells with Illumina high throughput sequencing. In a pre-analysis, samples were treated with high temperatures (fifteen minutes at 95 °C, **Figure 3A**) to distinguish viable cells (P2) from non-viable cells (P1) with flow cytometry. As a first pass at identifying sequences which influence translational efficiencies, we calculated the log fold enrichment of particular motifs in the no/low fluorescence FACS bin versus the high fluorescence bin across multiple independent libraries (**Figure 3B**). We found that while generally the same motifs had similar values across different libraries, there was some variation (e.g. while RPPG was depleted in the no/low fluorescence bin in the XXXX, XPXX, and XXPX libraries, it was slightly enriched in the XPPX library, **Figure 3B**). Furthermore, we also detected variation in mScarlet-I expression by microscopy among cells carrying the same reporter plasmid (**Supplementary Figure 3A-C**). Thus, the appearance of motifs in both FACS bins could reflect heterogeneity in plasmid-based fluorophore expression, a phenomenon also reporter in earlier studies [48,49] (**Supplementary Figure 3D**). Ultimately, analyzing the relationship between both bins provides the most precise measurement of translational efficiency. To validate our results, we compared the top twenty motifs with the lowest and the highest translational efficiencies with predicted ribosome pausing strengths associated with distinct motifs from earlier studies[21] (**Figure 3B**). We found good agreement between predicted pausing strengths and our results, indicating that our system is generally suitable for assessing translational efficiencies, despite variability in plasmid copy numbers within the bacterial population (**Supplementary Figure 3D**). Furthermore, while our analysis primarily focused translational efficiency at the amino acid level, it is likely that regulatory mechanisms at the codon level also play a role, as described in previous literature[41]. For example, several unique codon combinations were present for the amino acid motif RPPP in one of our libraries (**Supplementary Figure 4**). Differences in translational efficiency of different codon combinations which code for the same amino acid motif could also cause broad distributions of certain motifs across two distinct bins (**Figure 3B**).

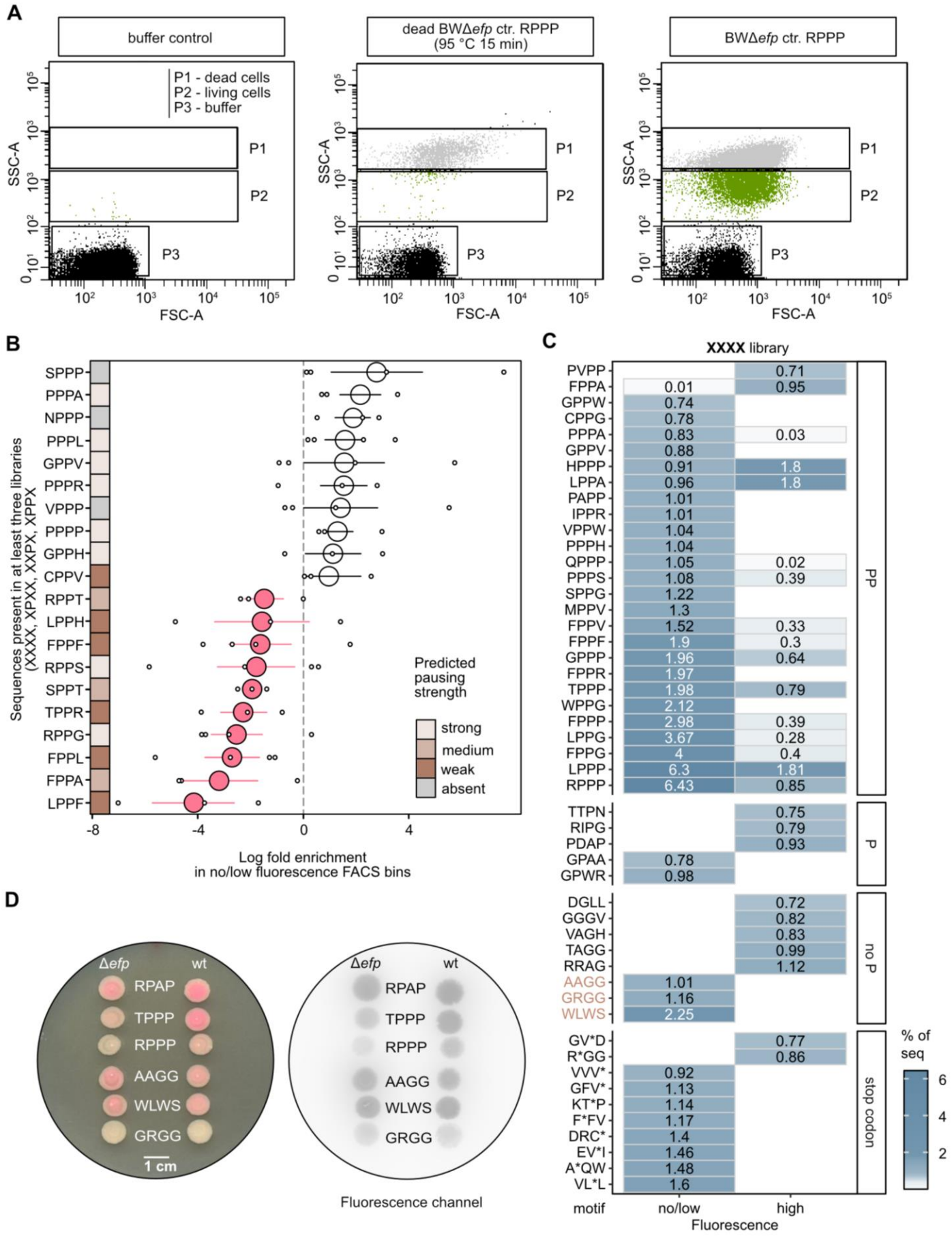


Figure 3. High throughput screening of four motif libraries (XPXX, XXPX, XPPX, XXXX).

(*see description on the next page*)

***Figure 3* (*description*).** **A** Gating strategy of living cells by flow cytometry. Living cells (P2) were sorted from the recorded buffer events (P3) and dead cells (P1) considering the Forward Scatter Area (FCS-A) and Side Scatter Area (SSC-A). **B** Summary of the motifs found in at least three out of four libraries (X-X-X-X; X-Pro-X-X; X-X-Pro-X; X-Pro-Pro-X) with the top ten highest and top ten lowest log fold enrichment. Libraries were expressed in *E. coli* BW25113 Δ*efp*. The log fold enrichment was calculated as the natural logarithm of the sequence count found in the no/low fluorescence FACS bin divided by the sequence count found in the high fluorescence FACS bin. Small circles- values in each library, big circles – mean values, lines – standard error of the mean (SEM). Big circles were coloured according to the calculated log fold enrichment value: <0, in pink; >0, in white. Pausing strength predictions were taken from Qi *et al*., 2018[21]. **C** The fifty most frequent amino acid motifs found in the XXXX motif library with the largest count differences between no/low fluorescence and high fluorescence FACS bins. The motifs were divided into categories of motifs containing at least two prolines (PP), one proline (P), no proline (no P) and a stop codon (stop codon, *). X represents all possible amino acids. **D** Translational efficiency measurements in Δ*efp* and wild-type (wt) cells carrying dual reporter plasmids with various motifs, analysed using a spot assay. Motifs analysed: RPAP (Arg-Pro-Ala-Pro), TPPP (Thr-Pro-Pro-Pro), RPPP (Arg-Pro-Pro-Pro), AAGG (Ala-Ala-Gly-Gly), WLWS (Trp-Leu-Trp-Ser) and GRGG (Gly-Arg-Gly-Gly). Right panel shows the fluorescence analysis of the spot assay plate.

A number of studies have reported on ribosomal pausing, which occurs at motifs containing consecutive prolines that can be rescued by EF-P[14-16,21,41]. So far only few studies report about motifs containing no prolines, that depend on EF-P for efficient translation[37,42]. Consequently, we wanted to see whether we could detect any non-proline motifs within the XXXX library that reduced translational efficiency of the fluorophore. We extracted the top fifty sequences with the largest difference in sequence counts between FACS bins and found four groups of motifs: those containing at least two prolines (PP), those containing one proline (P), those containing no proline (no P) and those containing a stop codon (**Figure 3C**). In accordance with our expectations, motifs containing a stop codon were highly abundant in the no/low fluorescence FACS bin (**Figure 3C**). In our system, stop codons cause the rapid termination of fluorophore translation, leading to the production of a non-functional protein lacking fluorescence. We found that certain motifs within the PP section appeared in both FACS bins but with differing abundance, which may be due to the heterogeneity phenomenon (**Supplementary Figure 3D**) and/or codon effects (**Supplementary Figure 4**). However, the majority of motifs in the PP section appeared in the no/low fluorescence FACS bins, consistent with previous studies indicating that the presence of at least two prolines hinders translation[16,21].

Interestingly, our experimental system identified two motifs containing a single proline (GPAA, GPWR) which decreased the efficiency of fluorophore translation. Additionally, three motifs without a proline (AAGG, GRGG, WLWS) were identified (**Figure 3C**). To verify these results, we transformed the Δ*efp* and wild-type strain with newly generated dual reporter plasmids containing the motifs AAGG, GRGG or WLWS, and spotted cells on agar assay plates. We found that these motifs negatively impact translational efficiency in the Δ*efp* mutant (**Figure 3D**), which is consistent with the observations shown in **Figure 3C**. The spot assay confirmed that the AAGG and WLWS motifs cause an intermediate reduction in translational efficiency of the fluorophore, while the GRGG motif causes a strong reduction in translation efficiency in the Δ*efp* mutant (**Figure 3D**). When EF-P and its modification machinery were present (wt), translation of the

fluorophore could be rescued for the control motifs TPPP and RPPP (**Figure 3D**). However, no rescue of fluorophore translation could be detected for the motifs AAGG, GRGG and WLWS (**Figure 3D**), suggesting that additional factors influence translation efficiency at these motifs.

Overall, we find that translational regulation appears to be more complex than previously thought. Consequently, our system offers a suitable platform to get more insights into this phenomenon in any laboratory.

Discussion

Translational coupling is a naturally occurring event in which the movement of ribosomes translating upstream genes also directs the translation of downstream genes. This mechanism ensures the coordinated and sequential synthesis of proteins involved in shared biochemical pathways or protein complexes, thereby ensuring efficient cellular function and effective resource allocation[50-54]. Translational coupling has been used in experimental studies to identify optimal translation initiation regions (TIRs) and enhance the synthesis of distinct proteins[43,44,55]. In the present study, we used translational coupling in the design of a dual reporter to investigate translational regulation. This system can be effectively adapted for use on agar assay plates, providing a dual output by selecting for fluorescence and viability, allowing uncomplicated screening.

As a component of the dual reporter system, mScarlet-I offer several advantages in plate-based translational efficiency screenings. Translational coupling of mScarlet-I to the resistance gene cassette (*cat*) resulted in the attenuation of *cat,* which improved reporter sensitivity (**Figure 1**). The presence of *cat* alone allowed only coarse distinctions in translational efficiencies — only motifs that are known to cause strong ribosome pausing and motifs that cause no pausing at all could be distinguished (**Supplementary Figure 1A-C**). This suggests that when present alone in the reporter, *cat* 'neutralises' chloramphenicol so effectively that small differences in ribosome pausing become indistinguishable. In contrast, the incorporation of mScarlet-I into the reporter allowed finer discrimination between pausing strengths caused by different motifs (**Figure 1**). In addition, the bright fluorescence, good colour contrast with the plate background (**Figure 1G** and **Figure 2B**) and ability to emit light in the visible red spectrum (FPbase ID: 6VVTK)[56,57] make mScarlet-I a perfect component of the reporter so that measurements could be performed without additional equipment.

Previous studies have demonstrated that in addition to consecutive prolines, sequences upstream and downstream of prolines can also influence translational efficiency in *E. coli*[16,38]*.* We were able to confirm and reproduce these observations, particularly in the context of downstream sequences (FPPX) in Δ*efp* (**Figure 2**). Motifs with no or intermediate influence on translational efficiency (FPPF, FPPI, FPPV and FPPY) and high influence on translational efficiency (FPPP)

(**Figure 2C**) were consistent with previously published proteome and ribosome profiling data[16,39]. However, we do not exclude the possibility that sequences further away from the motif, as well as the codon context of the motif, might also influence translational efficiency. This should be investigated in future studies.

Our investigations of motifs within the XXXX libraries, conducted using flow cytometry and high-throughput screening, identified new motifs containing one (GPAA, GPWR) or no prolines (AAGG, GRGG, WLWS) that affect translational efficiency (**Figure 3C and D**). Two of the motifs we identified were found to contain more than one glycine. Previous studies have demonstrated that sequences rich in glycine have the potential to cause ribosome pausing[37,58]. mRNA sequences that are similar to the Shine-Dalgarno (SD) sequence can trigger an SD-aSD (Shine-Dalgarno - anti-Shine-Dalgarno) interaction, which in turn causes ribosome pausing[5,59,60]. Nevertheless, it is still unclear how glycine codons affect translational pausing[61], as some studies suggest that internal SD sequences have no/low effect on translational speed[62,63]. The present study provides a platform that can support future research aimed at clarifying the influence of glycine on translational pausing in different sequence contexts.

In addition to EF-P, the ABCF ATPase YfmR has recently been shown to be involved in the translational rescue of polyproline motifs[64,65]. The reporter described here could easily be adapted to study the full spectrum of motifs rescued by this system in any laboratory. Additionally, while the spectrum of motifs that cause EF-P dependent ribosome pausing has been studied extensively, these experiments have primarily focused on *E. coli* and its β-lysylated post-translationally modified EF-P[16,21,39]. Our recent study showed differences in motif rescue preferences between wild-type *E. coli* and a mutant strain in which the native EF-P was replaced with an unmodified PGKGP subfamily EF-P[36]. This suggests that there is a subset of motifs that are preferentially rescued by different EF-P types. If so, different EF-P types could drive the evolution of bacterial genomes towards motifs that they preferentially rescue, particularly in highly expressed proteins where maximizing the translation rate is essential[21,66]. Such motif bias spectra could be studied with our reporter system.

Material and Methods

Bacterial growth

All bacterial strains used in this study were cultivated in lysogenic broth (LB) under agitation at 37°C. If required, growth media was solidified by the addition of 1.5 % (w/v) agar. The following antibiotics were used in this study: carbenicillin sodium salt (100 µg/ml), chloramphenicol (default concentration 34 µg/mL; reporter assays: 1.7 µg/mL, 2.3 µg/mL, 3.4 µg/mL and 3.8 µg/mL).

Molecular cloning

All strains, plasmids and DNA oligonucleotides used are listed in Supplementary Data 1. Kits, enzymes and polymerases were used according to the manufacturer's instructions. Q5® High-Fidelity DNA Polymerase (New England BioLab) and One*Taq*® DNA Polymerase were used for PCR amplification. DNA oligonucleotides were ordered from Merck. Amplified genes coding for the fluorophores (*sfgfp* and *mscarlet-I*) and the antibiotic resistance cassette (*chloramphenicol acetyltransferase*, CAT) were linked together with a hairpin loop (weak coupling)[43] by overlap PCR. DNA was isolated and purified with High-Yield PCR Cleanup and Gel Extraction Kit (Sued Laborbedarf Gauting). All DNA restriction digests were carried out using restriction enzymes in rCutSmart® buffer (New England BioLab). All ligations were performed with the T4 DNA ligase (New England BioLab) at 25 °C for 2 hours. Plasmids were isolated with the Hi Yield Plasmid Mini Kit (Sued Laborbedarf Gauting) and stored at -20°C. Plasmids were verified by colony PCR and sanger sequencing. All nucleotide sequences were analyzed with CLC Main Workbench version 8.1.2 (Qiagen).

Growth and fluorescence measurements in 96-well plates

To find the optimal concentration of chloramphenicol for the pausing strength measurements with the dual reporter, serial dilutions were tested. Plasmids pBAD24_motif_*sfgfp_cat* and pBAD24_motif_*mscarlet-I_cat* (motif: RPAP, TPPP or RPPP) were independently transformed into *E. coli* BW25113 Δ*efp* and incubated overnight at 37 °C on LB agar plates supplemented with 100 µg/ml carbenicillin sodium salt. Single colonies were inoculated into fresh LB supplemented with 100 µg/ml carbenicillin sodium salt and incubated overnight at 37°C under constant shaking (180 rpm). The overnight cultures were reinoculated into fresh LB supplemented with 100 µg/ml carbenicillin sodium salt and grown until reaching an absorption of 0.3 at 600 nm (OD_{600}). Each culture was transferred to a final OD_{600} of 0.01 into a 96-well plate, containing 0.2 % (w/v) arabinose and different concentrations of chloramphenicol (1.7 µg/mL, 2.3 µg/mL, 3.4 µg/mL, 3.8 µg/mL) with a final volume of 200 µL/well. Plates were incubated overnight at 37 °C under constant shaking (150 rpm). Endpoint OD_{600} was determined with Tecan Infinite 200 Pro (number of flashes: 25). Fluorescence was measured with GE Typhoon Trio Imager (laser: green, 532 nm; emission filter: 580 BP 30 Cy3; photo-multiplier tube, PMT: 450).

For all time-course fluorescence and growth measurements, the cells were grown for 20 hours in M9 minimal media[67] (composition: 33.7 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.55 mM NaCl, 9.35 mM NH_4Cl, 1 mM $MgSO_4$, 0.3 mM $CaCl_2$, 1 µg biotin, 1 µg thiamin, trace elements, 0.4 % [w/v] glucose) prior to the start of the measurements in the plate reader. Growth and fluorescence were monitored in 10 min intervals for 35 hours at 37 °C with Tecan Infinite 200 Pro (excitation wavelength: 565 nm; emission wavelength: 594 nm; gain: 50; number of flashes: 25; shaking between measurements: 180 rpm, orbital).

EF-P complementation assay plates were incubated for 24 to 34 hours until fluorescence measurements.

Generation of XPPX, PXXX, XPXX, XXPX, XXXP, XXXX libraries

The forward primers were designed to contain six to twelve degenerated nucleotides that code for all possible codons. The fixed prolines in the motif were coded by CCG codon. Degenerated codons were designed to appear fifteen nucleotides downstream of the start codon of the fluorophore gene. Libraries were amplified using the degenerated forward primer and reverse primer with complementary regions to the forward primers to enable plasmid circularization afterwards (**Figure 3A**). The PCR program used for library amplification was as follows: initial denaturation, 98 °C for 30 seconds (sec); amplification (35 cycles), 98 °C for 10 sec followed by 57°C for 30 sec and 72 °C for 180 sec; final extension, 72 °C for 120 sec. 40 µL of the PCR product was treated with *Dpn*I. Chemically competent *E. coli* MC1061 was transformed with the digested and purified PCR products to enable plasmid circularization by homologous recombination[43,44,47]. The transformation was transferred to a flask with 15 mL LB, supplemented with 100 µg/ml carbenicillin sodium salt, and incubated overnight at 37 °C under constant shaking (180 rpm). The libraries were isolated with the Hi Yield Plasmid Mini Kit (Sued Laborbedarf Gauting).

Screening of libraries on assay plates

The final libraries were prepared according to the protocol described by Rennig *et* al.[43] and Shilling *et* al.[44], with certain modifications. Chemically competent *E. coli* BW25113 Δ*efp* was transformed with 500 ng of the library with a 1-hour recovery phase in fresh LB without antibiotics. The transformants were inoculated into 3 mL fresh LB containing 100 µg/ml carbenicillin sodium salt and incubated overnight at 37 °C under constant shaking (180 rpm). 30 µL of this overnight culture was reinoculated into 3 mL fresh LB (containing 100 µg/ml carbenicillin sodium salt) and incubated for another round overnight at 37 °C under constant shaking (180 rpm). 100 µL of this overnight culture was inoculated into fresh 5 mL LB supplemented with 34 µg/mL chloramphenicol and grown until reaching an OD_{600} of 0.3. 10^3 - 10^5 cells were plated on assay agar plates, supplemented with 3.4 µg/mL chloramphenicol and 0.2 % (w/v) arabinose. Plates were incubated for 16 hours at 30 °C and kept afterwards for ~ 120 hours at 25°C. Fluorescence was measured with GE Typhoon Trio Imager (laser: green, 532 nm; emission filter: 580 BP 30 Cy3; pmt: 450).

Microscopy

To analyze the effect of different motifs on the production of mScarlet-I, chemically competent *E. coli* BW25113 Δ*efp* was transformed with the plasmids pBAD24_motif_*mscarlet-I_cat* (motif:

RPAP, TPPP or RPPP). Cells carrying the plasmids were cultivated overnight at 37 °C under constant shaking (180 rpm) in LB supplemented with 100 µg/ml carbenicillin sodium salt. The overnight cultures were inoculated into fresh LB supplemented with 100 µg/ml carbenicillin sodium salt and grown at 37 °C under constant shaking (180 rpm) until reaching an OD_{600} of 0.15. 0.2 % (w/v) arabinose was then added to the cultures and incubation continued for 24 hours. All cells were washed in phosphate buffered saline (PBS; 93.6 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 * 2 H_2O, 2 mM KH_2PO_4) and 2 µL of a culture with an OD_{600} of 0.5 were spotted on a 1 % (w/v) agarose pad (in PBS) and covered with a coverslip. Microscopic pictures were taken using the Leica DMi8 inverted microscope equipped with a Leica DFC365 FX camera (Wetzlar Germany). An excitation wavelength of 546 nm and a 605 nm emission filter with a 75-nm bandwidth were used for mScarlet-I fluorescence with an exposure of 20 ms, gain 1, and 100% intensity. To quantify the relative fluorescent intensities (RF) of single cells, phase contrast and fluorescent images were analyzed using the ImageJ[68] plugin MicrobeJ[69]. Default settings of MicrobeJ were used for cell segmentation (Fit shape, rod-shaped bacteria) apart from the following settings: area: 0.1-max µm^2; length: 1.2-5 µm; width: 0.1-1 µm; curvature 0.-0.15 and angularity 0.-0.25 for *E. coli* cells. In total 516 cells were quantified per strain. Background of the agarose pad was subtracted from each cell per field of view.

Flow cytometry and FACS

The influence of different motifs on the translational efficiency of the fluorophore mScarlet-I was analyzed using high throughput screening with flow cytometry and fluorescence associated cell sorting (FACS). Cells containing the dual reporter plasmid were grown in LB supplemented with 100 µg/ml carbenicillin sodium salt and 0.2 % (w/v) arabinose for 16 hours at 37 °C under constant shaking (180 rpm). The overnight incubation was necessary to provide enough time for mScarlet-I to mature before the FACS analysis. Cells were washed in PBS and seeded to an OD of 0.1 (10^8 cells) in 10 mL PBS (4°C) prior to the measurements. For the dead cells control, the culture was boiled at 95 °C for 15 min. The gating strategy for living cells is depicted in **Figure 4A**. Flow cytometry measurements and FACS were performed using FACSAria^{TM}FusionII Cell Sorter (BD Biosciences) with the following settings: 438 V(FSC), 277 V (SSC), 490 V (PE-Texas Red). Red fluorescence intensities were quantified using PE-Texas Red and displayed on a standard logarithmic scale. Flow cytometry data was processed with FACSDiva Software (BD Biosciences). Motifs found in cells with no/low fluorescence (intensity below $10^{2.1}$ on the logarithmic scale) were classified as those capable of slowing down the translational efficiency of the fluorophore. In contrast, motifs found in cells with high fluorescence (intensity above $10^{3.1}$ on the logarithmic scale) were classified as those having only a marginal influence on the translational efficiency of the fluorophore.

Library preparation for Illumina high throughput sequencing

Plasmid from each FACS bin were isolated with the Hi Yield Plasmid Mini Kit (Sued Laborbedarf) and served as template DNA for Illumina-specific amplification required for sequencing[70]. Custom made DNA oligos were designed (Merk), consisting of 21 bp Illumina primer sequence, distinct barcoding tag[70] and a sequence complementary to the dual reporter plasmid (forward primer: 81 bp upstream of the library motif, reverse primer: 167 bp downstream of the library motif; **Supplementary Figure S5**). Purified PCR products were loaded on a 1% (w/v) agarose gel to estimate the amount of each amplicon and pooled in equimolar concentrations[70]. The library was finalized for sequencing according to the manufacturer's protocols (http://supportres.illumina.com/documents/documentation/system_documentation/miseq/preparing-libraries-for-sequencing-on-miseq-15039740-d.pdf). Paired-end sequencing of 2 x 300 bp (base pairs) with two additional 8 bp index reads was performed on an Illumina MiSeq platform (Illumina Inc.) with v3 chemistry. All Sequencing data generated in this study is deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB77738

Bioinformatic analysis of Illumina library sequencing

Sequences were first quality filtered using FASTQ version 0.23.4[71]. Sequences were trimmed using a dynamic approach—if the mean quality of bases within a sliding window of 4 nucleotides dropped below a cutoff of 20, sequences were trimmed starting at the first base within the window (`--cut_right`). Paired reads were corrected using the overlapping region (`--correction`). Next, PANDAseq version 2.11 was used with default parameters to merge paired end sequences[72]. SABRE version 1.0 was used to demultiplex sequences[73], and custom python scripts were used to extract the sequence region containing the library. The R package Vegan version 2.6 was used to rarefy sequences to the level of the sample with the lowest sequence count (this level varied depending on the samples being compared but was never lower than 98335 sequences)[74]. All statistical analyses on the sequence data were done in R version 4.3.2[75], and all related figures were created with ggplot2 version 3.4.4[76].

Statistical analysis and reproducibility

All measurements are from at least three biological replicates except the plate-based library screening (**Figure 2C**, the count of found motifs is indicated in the figure) and FACS. Data comparisons were done using the student's unpaired two-sided t test (Microsoft Office Excel 2021) and are described in each figure legend. Error bars in bar graphs represent the standard deviation (SD). Values with a calculated p *value* below 0.05 were considered as significantly different.

Acknowledgements

We thank Dr. Andreas Brachmann and the Genomics Service Unit at the LMU for fruitful discussions and excellent support in the sequencing of the Illumina libraries. We thank the Core Facility Flow Cytometry at Biomedical Center Munich for the great technical support. We thank Lukas Voshagen for great technical assistance. This work was financially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant JU270/20-1, project number 449926427, and RTG2062 (Molecular Principles of Synthetic Biology) to K.J.

Author contributions

U.T. and K.J. designed the research; U.T. and K.B. performed the research; U.T., T.B. and S.B. analysed the data; U.T. and K.J. wrote the manuscript; K.J. funding acquisition.

Competing interests

The authors declare no competing interests.

References

1. Lu, J., Kobertz, W.R., and Deutsch, C. (2007). Mapping the electrostatic potential within the ribosomal exit tunnel. J Mol Biol *371*, 1378-1391. 10.1016/j.jmb.2007.06.038.
2. Lu, J., and Deutsch, C. (2008). Electrostatics in the ribosomal tunnel modulate chain elongation rates. J Mol Biol *384*, 73-86. 10.1016/j.jmb.2008.08.089.
3. Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat Struct Mol Biol *16*, 589-597. 10.1038/nsmb.1614.
4. de Smit, M.H., and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc Natl Acad Sci U S A *87*, 7668-7672. 10.1073/pnas.87.19.7668.
5. Wen, J.D., Lancaster, L., Hodges, C., Zeri, A.C., Yoshimura, S.H., Noller, H.F., Bustamante, C., and Tinoco, I. (2008). Following translation by single ribosomes one codon at a time. Nature *452*, 598-603. 10.1038/nature06716.
6. Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. Science *324*, 255-258. 10.1126/science.1170160.
7. Chen, C., Zhang, H., Broitman, S.L., Reiche, M., Farrell, I., Cooperman, B.S., and Goldman, Y.E. (2013). Dynamics of translation by single ribosomes through mRNA secondary structures. Nat Struct Mol Biol *20*, 582-588. 10.1038/nsmb.2544.
8. Varenne, S., Buc, J., Lloubes, R., and Lazdunski, C. (1984). Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains. J Mol Biol *180*, 549-576. 10.1016/0022-2836(84)90027-5.
9. Nedialkova, D.D., and Leidel, S.A. (2015). Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. Cell *161*, 1606-1618. 10.1016/j.cell.2015.05.022.
10. Liu, F., Clark, W., Luo, G., Wang, X., Fu, Y., Wei, J., Wang, X., Hao, Z., Dai, Q., Zheng, G., et al. (2016). ALKBH1-Mediated tRNA Demethylation Regulates Translation. Cell *167*, 816-828 e816. 10.1016/j.cell.2016.09.038.
11. Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol *2*, 13-34. 10.1093/oxfordjournals.molbev.a040335.
12. Komar, A.A. (2016). The Yin and Yang of codon usage. Hum Mol Genet *25*, R77-R85. 10.1093/hmg/ddw207.
13. Tanner, D.R., Cariello, D.A., Woolstenhulme, C.J., Broadbent, M.A., and Buskirk, A.R. (2009). Genetic identification of nascent peptides that induce ribosome stalling. J Biol Chem *284*, 34809-34818. 10.1074/jbc.M109.039040.
14. Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., and Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science *339*, 82-85. 10.1126/science.1228985.
15. Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85-88. 10.1126/science.1229017.
16. Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumae, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., and Wilson, D.N. (2013). Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. Proc Natl Acad Sci U S A *110*, 15265-15270. 10.1073/pnas.1310642110.
17. Muto, H., and Ito, K. (2008). Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin. Biochem Biophys Res Commun *366*, 1043-1047. 10.1016/j.bbrc.2007.12.072.
18. Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M.V. (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. J Biol Chem *283*, 32229-32235. 10.1074/jbc.M805316200.
19. Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M., and Forster, A.C. (2009). Slow peptide bond formation by proline and other N-alkylamino acids in translation. Proc Natl Acad Sci U S A *106*, 50-54. 10.1073/pnas.0809211106.
20. Doerfel, L.K., Wohlgemuth, I., Kubyshkin, V., Starosta, A.L., Wilson, D.N., Budisa, N., and Rodnina, M.V. (2015). Entropic Contribution of Elongation Factor P to Proline Positioning

at the Catalytic Center of the Ribosome. J Am Chem Soc *137*, 12997-13006. 10.1021/jacs.5b07427.

21. Qi, F., Motz, M., Jung, K., Lassak, J., and Frishman, D. (2018). Evolutionary analysis of polyproline motifs in *Escherichia coli* reveals their regulatory role in translation. PLoS Comput Biol *14*, e1005987. 10.1371/journal.pcbi.1005987.
22. Pinheiro, B., Scheidler, C.M., Kielkowski, P., Schmid, M., Forne, I., Ye, S., Reiling, N., Takano, E., Imhof, A., Sieber, S.A., et al. (2020). Structure and Function of an Elongation Factor P Subfamily in Actinobacteria. Cell Rep *30*, 4332-4342 e4335. 10.1016/j.celrep.2020.03.009.
23. Morgan, A.A., and Rubenstein, E. (2013). Proline: the distribution, frequency, positioning, and common functional roles of proline and polyproline sequences in the human proteome. PLoS One *8*, e53785. 10.1371/journal.pone.0053785.
24. Rath, A., Davidson, A.R., and Deber, C.M. (2005). The structure of "unstructured" regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. Biopolymers *80*, 179-185. 10.1002/bip.20227.
25. Adzhubei, A.A., Sternberg, M.J., and Makarov, A.A. (2013). Polyproline-II helix in proteins: structure and function. J Mol Biol *425*, 2100-2132. 10.1016/j.jmb.2013.03.018.
26. Motz, M., and Jung, K. (2018). The role of polyproline motifs in the histidine kinase EnvZ. PLoS One *13*, e0199782. 10.1371/journal.pone.0199782.
27. Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Mol Cell *51*, 35-45. 10.1016/j.molcel.2013.04.021.
28. Behshad, E., Ruzicka, F.J., Mansoorabadi, S.O., Chen, D., Reed, G.H., and Frey, P.A. (2006). Enantiomeric free radicals and enzymatic control of stereochemistry in a radical mechanism: the case of lysine 2,3-aminomutases. Biochemistry *45*, 12639-12646. 10.1021/bi061328t.
29. Yanagisawa, T., Sumida, T., Ishii, R., Takemoto, C., and Yokoyama, S. (2010). A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. Nat Struct Mol Biol *17*, 1136-1143. 10.1038/nsmb.1889.
30. Navarre, W.W., Zou, S.B., Roy, H., Xie, J.L., Savchenko, A., Singer, A., Edvokimova, E., Prost, L.R., Kumar, R., Ibba, M., and Fang, F.C. (2010). PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in *Salmonella enterica*. Mol Cell *39*, 209-221. 10.1016/j.molcel.2010.06.021.
31. Peil, L., Starosta, A.L., Virumae, K., Atkinson, G.C., Tenson, T., Remme, J., and Wilson, D.N. (2012). Lys34 of translation elongation factor EF-P is hydroxylated by YfcM. Nat Chem Biol *8*, 695-697. 10.1038/nchembio.1001.
32. Lassak, J., Keilhauer, E.C., Furst, M., Wuichet, K., Godeke, J., Starosta, A.L., Chen, J.M., Sogaard-Andersen, L., Rohr, J., Wilson, D.N., et al. (2015). Arginine-rhamnosylation as new strategy to activate translation elongation factor P. Nat Chem Biol *11*, 266-270. 10.1038/nchembio.1751.
33. Rajkovic, A., Erickson, S., Witzky, A., Branson, O.E., Seo, J., Gafken, P.R., Frietas, M.A., Whitelegge, J.P., Faull, K.F., Navarre, W., et al. (2015). Cyclic Rhamnosylated Elongation Factor P Establishes Antibiotic Resistance in *Pseudomonas aeruginosa*. mBio *6*, e00823. 10.1128/mBio.00823-15.
34. Rajkovic, A., Hummels, K.R., Witzky, A., Erickson, S., Gafken, P.R., Whitelegge, J.P., Faull, K.F., Kearns, D.B., and Ibba, M. (2016). Translation Control of Swarming Proficiency in *Bacillus subtilis* by 5-Amino-pentanolylated Elongation Factor P. J Biol Chem *291*, 10976-10985. 10.1074/jbc.M115.712091.
35. Hummels, K.R., Witzky, A., Rajkovic, A., Tollerson, R., 2nd, Jones, L.A., Ibba, M., and Kearns, D.B. (2017). Carbonyl reduction by YmfI in *Bacillus subtilis* prevents accumulation of an inhibitory EF-P modification state. Mol Microbiol *106*, 236-251. 10.1111/mmi.13760.
36. Tomasiunaite, U., Kielkowski, P., Krafczyk, R., Forne, I., Imhof, A., and Jung, K. (2024). Decrypting the functional design of unmodified translation elongation factor P. Cell Rep *43*, 114063. 10.1016/j.celrep.2024.114063.
37. Hersch, S.J., Wang, M., Zou, S.B., Moon, K.M., Foster, L.J., Ibba, M., and Navarre, W.W. (2013). Divergent protein motifs direct elongation factor P-mediated translational regulation in *Salmonella enterica* and *Escherichia coli*. mBio *4*, e00180-00113. 10.1128/mBio.00180-13.

38. Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Virumae, K., Tenson, T., Remme, J., Jung, K., and Wilson, D.N. (2014). Translational stalling at polyproline stretches is modulated by the sequence context upstream of the stall site. Nucleic Acids Res *42*, 10711-10719. 10.1093/nar/gku768.
39. Woolstenhulme, C.J., Guydosh, N.R., Green, R., and Buskirk, A.R. (2015). High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. Cell Rep *11*, 13-21. 10.1016/j.celrep.2015.03.014.
40. Pfab, M., Kielkowski, P., Krafczyk, R., Volkwein, W., Sieber, S.A., Lassak, J., and Jung, K. (2021). Synthetic post-translational modifications of elongation factor P using the ligase EpmA. FEBS J *288*, 663-677. 10.1111/febs.15346.
41. Krafczyk, R., Qi, F., Sieber, A., Mehler, J., Jung, K., Frishman, D., and Lassak, J. (2021). Proline codon pair selection determines ribosome pausing strength and translation efficiency in bacteria. Commun Biol *4*, 589. 10.1038/s42003-021-02115-z.
42. Sieber, A., Parr, M., von Ehr, J., Dhamotharan, K., Kielkowski, P., Brewer, T., Schäpers, A., Krafczyk, R., Qi, F., Schlundt, A., et al. (2024). EF-P and its paralog EfpL (YeiP) differentially control translation of proline-containing sequences. bioRxiv, 2024.2004.2015.589488. 10.1101/2024.04.15.589488.
43. Rennig, M., Martinez, V., Mirzadeh, K., Dunas, F., Rojsater, B., Daley, D.O., and Norholm, M.H.H. (2018). TARSyn: Tunable Antibiotic Resistance Devices Enabling Bacterial Synthetic Evolution and Protein Production. ACS Synth Biol *7*, 432-442. 10.1021/acssynbio.7b00200.
44. Shilling, P.J., Mirzadeh, K., Cumming, A.J., Widesheim, M., Kock, Z., and Daley, D.O. (2020). Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*. Commun Biol *3*, 214. 10.1038/s42003-020-0939-8.
45. Verma, M., Choi, J., Cottrell, K.A., Lavagnino, Z., Thomas, E.N., Pavlovic-Djuranovic, S., Szczesny, P., Piston, D.W., Zaher, H.S., Puglisi, J.D., and Djuranovic, S. (2019). A short translational ramp determines the efficiency of protein synthesis. Nat Commun *10*, 5774. 10.1038/s41467-019-13810-1.
46. Kolbe, K., Bell, A.C., Prosser, G.A., Assmann, M., Yang, H.J., Forbes, H.E., Gallucci, S., Mayer-Barber, K.D., Boshoff, H.I., and Barry Iii, C.E. (2020). Development and Optimization of Chromosomally-Integrated Fluorescent *Mycobacterium tuberculosis* Reporter Constructs. Front Microbiol *11*, 591866. 10.3389/fmicb.2020.591866.
47. Mirzadeh, K., Martinez, V., Toddo, S., Guntur, S., Herrgard, M.J., Elofsson, A., Norholm, M.H., and Daley, D.O. (2015). Enhanced Protein Production in *Escherichia coli* by Optimization of Cloning Scars at the Vector-Coding Sequence Junction. ACS Synth Biol *4*, 959-965. 10.1021/acssynbio.5b00033.
48. Jahn, M., Vorpahl, C., Hubschmann, T., Harms, H., and Muller, S. (2016). Copy number variability of expression plasmids determined by cell sorting and Droplet Digital PCR. Microb Cell Fact *15*, 211. 10.1186/s12934-016-0610-8.
49. Boy, C., Lesage, J., Alfenore, S., Gorret, N., and Guillouet, S.E. (2020). Plasmid expression level heterogeneity monitoring via heterologous eGFP production at the single-cell level in *Cupriavidus necator*. Appl Microbiol Biotechnol *104*, 5899-5914. 10.1007/s00253-020-10616-w.
50. Aksoy, S., Squires, C.L., and Squires, C. (1984). Translational coupling of the trpB and trpA genes in the *Escherichia coli* tryptophan operon. J Bacteriol *157*, 363-367. 10.1128/jb.157.2.363-367.1984.
51. de Smit, M.H., and van Duin, J. (1990). Control of prokaryotic translational initiation by mRNA secondary structure. Prog Nucleic Acid Res Mol Biol *38*, 1-35. 10.1016/s0079-6603(08)60707-2.
52. Lesage, P., Chiaruttini, C., Graffe, M., Dondon, J., Milet, M., and Springer, M. (1992). Messenger RNA secondary structure and translational coupling in the *Escherichia coli* operon encoding translation initiation factor IF3 and the ribosomal proteins, L35 and L20. J Mol Biol *228*, 366-386. 10.1016/0022-2836(92)90827-7.
53. Rex, G., Surin, B., Besse, G., Schneppe, B., and McCarthy, J.E. (1994). The mechanism of translational coupling in *Escherichia coli*. Higher order structure in the atpHA mRNA acts as a conformational switch regulating the access of de novo initiating ribosomes. J Biol Chem *269*, 18118-18127.

54. Chiaruttini, C., Milet, M., and Springer, M. (1997). Translational coupling by modulation of feedback repression in the IF3 operon of *Escherichia coli*. Proc Natl Acad Sci U S A *94*, 9208-9213. 10.1073/pnas.94.17.9208.
55. Ferro, R., Rennig, M., Hernandez-Rollan, C., Daley, D.O., and Norholm, M.H.H. (2018). A synbio approach for selection of highly expressed gene variants in Gram-positive bacteria. Microb Cell Fact *17*, 37. 10.1186/s12934-018-0886-y.
56. Lambert, T.J. (2019). FPbase: a community-editable fluorescent protein database. Nat Methods *16*, 277-278. 10.1038/s41592-019-0352-8.
57. Bindels, D.S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K.E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M.A., and Gadella, T.W., Jr. (2017). mScarlet: a bright monomeric red fluorescent protein for cellular imaging. Nat Methods *14*, 53-56. 10.1038/nmeth.4074.
58. Woolstenhulme, C.J., Parajuli, S., Healey, D.W., Valverde, D.P., Petersen, E.N., Starosta, A.L., Guydosh, N.R., Johnson, W.E., Wilson, D.N., and Buskirk, A.R. (2013). Nascent peptides that block protein synthesis in bacteria. Proc Natl Acad Sci U S A *110*, E878-887. 10.1073/pnas.1219536110.
59. Li, G.W., Oh, E., and Weissman, J.S. (2012). The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature *484*, 538-541. 10.1038/nature10965.
60. Chevance, F.F., Le Guyon, S., and Hughes, K.T. (2014). The effects of codon context on in vivo translation speed. PLoS Genet *10*, e1004392. 10.1371/journal.pgen.1004392.
61. Rodnina, M.V. (2016). The ribosome in action: Tuning of translational efficiency and protein folding. Protein Sci *25*, 1390-1406. 10.1002/pro.2950.
62. Mohammad, F., Woolstenhulme, C.J., Green, R., and Buskirk, A.R. (2016). Clarifying the Translational Pausing Landscape in Bacteria by Ribosome Profiling. Cell Rep *14*, 686-694. 10.1016/j.celrep.2015.12.073.
63. Borg, A., and Ehrenberg, M. (2015). Determinants of the rate of mRNA translocation in bacterial protein synthesis. J Mol Biol *427*, 1835-1847. 10.1016/j.jmb.2014.10.027.
64. Hong, H.R., Prince, C.R., Tetreault, D.D., Wu, L., and Feaga, H.A. (2024). YfmR is a translation factor that prevents ribosome stalling and cell death in the absence of EF-P. Proc Natl Acad Sci U S A *121*, e2314437121. 10.1073/pnas.2314437121.
65. Takada, H., Fujiwara, K., Atkinson, G.C., Chiba, S., and Hauryliuk, V. (2024). Resolution of ribosomal stalling by EF-P and ABCF ATPases YfmR and YkpA/YbiT. Nucleic Acids Res. 10.1093/nar/gkae556.
66. Brewer, T.E., and Wagner, A. (2022). Translation stalling proline motifs are enriched in slow-growing, thermophilic, and multicellular bacteria. ISME J *16*, 1065-1073. 10.1038/s41396-021-01154-y.
67. Miller, J.H. (1972). Experiments in molecular genetics (Cold Spring Harbor, N.Y., [Cold Spring Harbor, N.Y.]).
68. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675. 10.1038/nmeth.2089.
69. Ducret, A., Quardokus, E.M., and Brun, Y.V. (2016). MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. Nat Microbiol *1*, 16077. 10.1038/nmicrobiol.2016.77.
70. Unterseher, M., Siddique, A.B., Brachmann, A., and Persoh, D. (2016). Diversity and Composition of the Leaf Mycobiome of Beech (*Fagus sylvatica*) Are Affected by Local Habitat Conditions and Leaf Biochemistry. PLoS One *11*, e0152878. 10.1371/journal.pone.0152878.
71. Chen, S. (2023). Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. iMeta *2*, e107. https://doi.org/10.1002/imt2.107.
72. Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G., and Neufeld, J.D. (2012). PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics *13*, 31. 10.1186/1471-2105-13-31.
73. Joshi, N. (2011). sabre - A barcode demultiplexing and trimming tool for FastQ files. https://github.com/najoshi/sabre.
74. Oksanen J, S.G., Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux

B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2022). Community ecology Package. vegan: Community Ecology Package .R package version 2.6-4. https://CRAN.R-project.org/package=vegan.

75. R Core Team. R: A Language and Environment for Statistical Computing. (2021). https://www.R-project.org/.

76. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://CRAN.R-project.org/package=ggplot2.